# nature research

Corresponding author(s): Albert Cheng

Last updated by author(s): 5/21/2020

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist .

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☒ | ☐ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒ | ☐ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☒ | ☐ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☒ | ☐ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ | For null hypothesis testing, the test statistic (e.g. $F$, $t$, $r$) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| Data collection | qRT-PCR measurements were collected with LightCycler® 480 Software (Roche) |
|---|---|
| Data analysis | FASTQ raw sequence files were quality checked with FASTQC version 0.11.5. Low quality reads and adapter sequences were trimmed by Trimmomatic version 0.33. Mapping was performed by STAR aligner version 2.5.3 to the human genome (USCS RefSeq hg38 annotation). Transcriptome-wide splicing events were identified by rMATS version 3.2.5 using UCSC RefSeq hg38 GTF file as annotation. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are included in the published article and in the Supplementary Information, and are available from the corresponding author upon reasonable request. The source data underlying all the figures and supplementary figures, as well as the qRT-PCR plot data are included as Source Data. Data containing RNA-seq raw sequencing read files were deposited onto sequencing read archive (SRA) with accession number PRJNA624911. Plasmids are available on the Addgene repository at https://www.addgene.org/browse/article/28196786/

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

[✗] Life sciences          [ ] Behavioural & social sciences          [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | No sample size calculation was performed. For qRT-PCR, a field-standard of 3 is used. |
| Data exclusions | No data were excluded |
| Replication | Experiments were replicated at least once. For qRT-PCR, triplicates were performed on top of biological replication |
| Randomization | Samples were not allocated into groups, thus randomization is not relevant for this study. |
| Blinding | Blinding is not relevant for this study since there was no group allocation |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| [✗] | [ ] Antibodies |
| [ ] | [✗] Eukaryotic cell lines |
| [✗] | [ ] Palaeontology |
| [✗] | [ ] Animals and other organisms |
| [✗] | [ ] Human research participants |
| [✗] | [ ] Clinical data |

## Methods

| n/a | Involved in the study |
|---|---|
| [✗] | [ ] ChIP-seq |
| [✗] | [ ] Flow cytometry |
| [✗] | [ ] MRI-based neuroimaging |

# Eukaryotic cell lines

Policy information about cell lines

| | |
|---|---|
| Cell line source(s) | HEK293T, U2OS, HeLa (ATCC), GM03813 (Coriell institute) |
| Authentication | None of the cell lines used were authenticated |
| Mycoplasma contamination | Cell lines were not tested for mycoplasma contaimination |
| Commonly misidentified lines (See ICLAC register) | No commonly misidentified cell lines were used in the study. |